# Transparency, Fairness, Data Protection, Neutrality: Data Management Challenges in the Face of New Regulation

Serge Abiteboul, Julia Stoyanovich

# Transparency, Fairness, Data Protection, Neutrality: Data Management Challenges in the Face of New Regulation

SERGE ABITEBOUL, Inria & Ecole Normale Supérieure, France
JULIA STOYANOVICH, New York University, USA

The data revolution continues to transform every sector of science, industry and government. Due to the incredible impact of data-driven technology on society, we are becoming increasingly aware of the imperative to use data and algorithms responsibly — in accordance with laws and ethical norms. In this article we discuss three recent regulatory frameworks: the European Union's General Data Protection Regulation (GDPR), the New York City Automated Decisions Systems (ADS) Law, and the Net Neutrality principle, that aim to protect the rights of individuals who are impacted by data collection and analysis. These frameworks are prominent examples of a global trend: Governments are starting to recognize the need to regulate data-driven algorithmic technology.

Our goal in this paper is to bring these regulatory frameworks to the attention of the data management community, and to underscore the technical challenges they raise and which we, as a community, are well-equipped to address. The main take-away of this article is that legal and ethical norms cannot be incorporated into data-driven systems as an afterthought. Rather, we must think in terms of responsibility by design, viewing it as a systems requirement.

CCS Concepts: • **Information systems** → **Data management systems**; • **Social and professional topics** → **Computing / technology policy**; *Technology audits*; • **Applied computing** → *Law*;

Additional Key Words and Phrases: transparency; fairness; data protection; neutrality; responsible data science

**ACM Reference Format:**
Serge Abiteboul and Julia Stoyanovich. 2018. Transparency, Fairness, Data Protection, Neutrality: Data Management Challenges in the Face of New Regulation. *ACM J. Data Inform. Quality* 0, 0, Article 0 ( 2018), 9 pages. https://doi.org/0000001.0000001

## 1 INTRODUCTION

The data revolution continues to transform every sector of science, industry and government. Due to the incredible impact of data-driven technology on society, we are becoming increasingly aware of the imperative to use data and algorithms *responsibly* — in accordance with laws and ethical norms. The goal of this article is to underscore the technical challenges raised by recent legal and regulatory frameworks, which the data management community is well-equipped to address.

We discuss three recent frameworks: the European Union's General Data Protection Regulation (GDPR) [The European Union 2016], the New York City Automated Decisions Systems (ADS) Law [The New York City Council 2017], and the Net Neutrality principle. These frameworks are prominent examples of a global trend: Governments are starting to recognize the need to regulate data-driven algorithmic technology. The GDPR and the NYC ADS Law aim to protect

Authors' addresses: Serge Abiteboul, Inria & Ecole Normale Supérieure, France, Serge.Abiteboul@inria.fr; Julia Stoyanovich, New York University, USA, stoyanovich@nyu.edu.


Manuscript submitted to ACM

the rights of individuals who are impacted by data collection and analysis, while the Net Neutrality principle ensures that services are being treated equitably. Yet, despite the focus on organizations, rights of individuals also figure prominently in the neutrality debate: One of the imperatives is that individuals should be able to enjoy freedom of choice and expression on-line. We will give some legal context on neutrality by discussing the EU Regulation 2015/2120 [The European Parliament AND Council 2015], the Indian Net Neutrality Regulatory Framework [Government of India, Ministry 2018], and the ongoing regulatory debate on Net Neutrality in the US.

Our goal in this paper is to bring these regulatory frameworks to the attention of the data management community. The main take-away of this article is that legal norms cannot be incorporated into data-driven systems as an afterthought. Rather, we must think in terms of *responsibility by design*, viewing it as a systems requirement.

### 1.1 The General Data Protection Regulation

The European Union recently enacted a sweeping regulatory framework known as the General Data Protection Regulation, or the GDPR [The European Union 2016]. The regulation was adopted in April 2016, and became enforceable about two years later, on May 25, 2018. The GDPR aims to protect the rights and freedoms of natural persons with regard to how their personal data is processed, moved and exchanged (Article 1). The GDPR is broad in scope, and applies to "the processing of personal data wholly or partly by automated means" (Article 2), both in the private sector and in the public sector. Personal data is broadly construed, and refers to any information relating to an identified or identifiable natural person, called the *data subject* (Article 4). In this article we focus on the following salient points of the regulation:

- lawful processing of data is predicated on the data subject's informed consent, stating whether their personal data can be used, and for what purpose (Articles 6, 7);
- the data subject has a right to correct any errors in their data ("right to rectification", Article 16), to withdraw their data from the system ("right to erasure", Article 17), and to move data from one data processor to another ("right to portability", Article 20);
- the data subject has the right to be informed about the collection and use of their data. [1]

The primary focus of the GDPR is on protecting the rights of data subjects, by giving them insight into, and control over, the collection and processing of their personal data. Providing insight, in response to the "right to be informed", requires technical methods for algorithmic and data transparency, which we will discuss in Section 2. We will also discuss the challenges inherent in giving individuals an ability to erase or move their data in Section 4.

### 1.2 The New York City Algorithmic Decision Systems Law

New York City recently passed a law [The New York City Council 2017] requiring that a task force be put in place to survey the current use of "automated decision systems" (ADS), defined as "computerized implementations of algorithms, including those derived from machine learning or other data processing or artificial intelligence techniques, which are used to make or assist in making decisions," in City agencies. The task force is working to develop a set of recommendations for enacting algorithmic transparency by the agencies, and will propose procedures for:

- requesting and receiving an explanation of an algorithmic decision affecting an individual (Section 3 (b));

[1] https://gdpr-info.eu/issues/right-to-be-informed/

- interrogating automated decision systems for bias and discrimination against members of legally protected groups, and addressing instances in which a person is harmed based on membership in such groups (Sections 3 (c) and (d));
- assessing how automated decision systems function and are used, and archiving the systems together with the data they use (Sections 3 (e) and (f)).

In contrast to the GDPR, which is very broad in scope, the NYC ADS Law only regulates City agencies in their use of algorithms and data, and does not directly apply to private companies. However, because government agencies often procure systems and components from industry partners, the Law will likely impact industry practices. Further, while New York is the first US city to pass a law of this kind, we expect other US municipalities to follow with similar legal frameworks or recommendations in the near future.

The primary focus of the NYC ADS Law is on algorithmic transparency, which, in turn, cannot be achieved without data transparency [Stoyanovich and Howe 2018]. As we discussed in Section 1.1, transparency is also an implicit requirement of the GDPR, stemming from the "right to be informed". We will discuss the role that the data management community can play in enabling data transparency in Section 2.

The NYC ADS Law further requires fair and equitable treatment of individuals, mandating that ADS safeguard against bias and discrimination, and provide transparency in this regard. We will discuss fairness in Section 3, and will propose some research directions for the data management community that are complementary to the rich and rapidly expanding body of work on fairness in machine learning.

### 1.3 The Net Neutrality Principle

Net Neutrality is the principle that Internet Service Providers (ISPs) should not discriminate or charge differently based on the message source (the content provider), its destination (the user), or its content. The concept was articulated by Tim Wu in 2003 [Wu 2003].

According to Net Neutrality, an ISP cannot block or throttle video streams from YouTube (negative discrimination), or enable free access to Facebook out of package (a kind of positive discrimination). A September 2018 report from Northeastern University and the University of Massachusetts, Amherst, found that US telecommunications companies are indeed slowing internet traffic to and from those two sites in particular, along with other popular apps [Kharif 2018; Molavi Kakhki et al. 2015]. Of course, there are limits to the non-discrimination, such as blocking pornographic material for young Internet users, filtering hate speech in some countries, or guaranteeing quality for emergency services.

In the European Union, Net Neutrality is guaranteed by EU Regulation 2015/2120 [The European Parliament AND Council 2015], although different countries may interpret the regulation differently. For example, some forms of zero-rating, the practice of providing Internet access without financial cost as a means of positive discrimination, are legal in some EU countries but not in others. Since 2018, India has perhaps the world's strongest Net Neutrality rules [Government of India, Ministry of Communicatior 2018]. In general, more and more countries are adopting Net Neutrality regulations, with a notable exception. In the United States, the Federal Communications Commission (FCC) issued its Open Internet Order in 2015, reclassifying Internet access — previously classified as an information service — as a common carrier telecommunications service, thereby enforcing some form of Net Neutrality. However, in 2017, under the chairmanship of Ajit Pai, the FCC officially repealed Net Neutrality rules.

## 2 ALGORITHMIC AND DATA TRANSPARENCY

ProPublica's story on "machine bias" in an algorithm used for sentencing defendants [Angwin et al. 2016] amplified calls to make algorithms more transparent and accountable [Kroll et al. 2017]. Transparency and accountability are intrinsically linked with trust, and are of particular importance when algorithmic systems are integrated into government processes, assisting humans in their decision-making tasks, and sometimes even replacing humans. Transparency of government is a core democratic value, which compels us to develop technological solutions that both increase government efficiency and can be made transparent to the public.

A narrow interpretation of algorithmic transparency requires that the source code of a system be made publicly available. This is a significant step towards transparency (as long as the posted code is readable, well-documented and complete), but it is rarely sufficient. One of the reasons for this, of particular relevance to the data management community, is that meaningful transparency of algorithmic processes cannot be achieved without transparency of data [Stoyanovich and Howe 2018].

What is data transparency, and how can we achieve it? One immediate interpretation of this term in the context of predictive analytics includes "making the training and validation datasets publicly available." However, while data should be made open whenever possible, much of it is sensitive and cannot be shared directly. That is, data transparency is in tension with the privacy of individuals who are included in the dataset. In light of this, we may adopt the following alternative interpretation of data transparency: In addition to releasing training and validation datasets whenever possible, vendors should make publicly available summaries of relevant statistical properties of the datasets that can aid in interpreting the decisions made using this data, while applying state-of-the-art methods to preserve the privacy of individuals (such as differential privacy [Dwork and Roth 2014]). When appropriate, privacy-preserving synthetic datasets can be released in lieu of real datasets to expose certain features of the data [Ping et al. 2017].

An important aspect of data transparency is interpretability — surfacing the statistical properties of a dataset, the methodology that was used to produce it, and, ultimately, substantiating its "fitness for use" in the context of a specific automated decision system or task. This consideration of a specific use is particularly important because datasets are increasingly used outside the original context for which they were intended. The data management community can begin addressing these challenges by building on the significant body of work on data profiling (see [Abedjan et al. 2017] for a recent tutorial), with an eye on the new legal requirements.

Interpretability rests on making explicit the interactions between the program and the data on which it acts. This property is important both when an automated decision system is interrogated for systematic bias and discrimination, and when it is asked to explain an algorithmic decision that affects an individual. For example, suppose that a system scores and ranks individuals for access to a service. If an individual enters her data and receives the result — say, a score of 42 — this number alone provides no information about why she was scored in this way, how she compares to others, and what she can do to potentially improve her outcome. A prominent example of a system of this kind, which is both opaque and extremely impactful, is the FICO credit scoring system in the US [Citron and Pasquale 2014].

The data management research community is well-positioned to contribute to developing new methods for interpretability. These new contributions can naturally build on a rich body of work on data provenance (see [Herschel et al. 2017] for a recent survey), on recent work on explaining classifiers [Ribeiro et al. 2016] and auditing black box models using causal framework [Datta et al. 2016], and on automatically generating "nutritional labels" for data and models [Yang et al. 2018].

## 3 FAIRNESS

We can all agree that algorithmic decision-making should be fair, even if we do not agree on the definition of fairness. But isn't this about algorithm design? Why is this a data problem? Indeed, the machine learning and data mining research communities are actively working on methods for enabling fairness of specific algorithms and their outputs, with a particular focus on classification problems (see, for example, [Dwork et al. 2012; Feldman et al. 2015; Friedler et al. 2016; Hajian and Domingo-Ferrer 2013; Kamiran et al. 2013; Kleinberg et al. 2017; Romei and Ruggieri 2014] and proceedings of the recently established ACM Conference on Fairness, Accountability, and Transparency (ACM FAT*) [2]). While important, these approaches focus solely on the final step in the data science lifecycle, and are thus limited by the assumption that input datasets are clean and reliable.

Data-driven algorithmic decision making usually requires multiple pre-processing stages to address messy input and render it ready for analysis [Jagadish et al. 2014]. This pre-processing, which includes data cleaning, integration, querying and ranking, is often the source of algorithmic bias [Kirkpatrick 2017; Stoyanovich et al. 2017], and so reasoning about sources of bias, and mitigating unfairness upstream from the final step of data analysis, is potentially more impactful.

For example, much research goes into ensuring statistical parity — a requirement that the demographics of those receiving a particular outcome, (e.g., a positive or negative classification), are identical to the demographics of the population as a whole. Suppose that the input to a binary classifier contains 900 men and 100 women, but that it is known that women represent 50% of the over-all population, and so achieving statistical parity amounts to enforcing a 50-50 gender balance among the positively classified individuals. That is, all else being equal, a woman in the input to the classifier is far more likely to receive a positive classification than a man. An alternative is to observe the following: If the input to the classifier was produced by a SQL query, and if relaxing the query would make the input more balanced (e.g., 1000 men and 500 women), then a more effective way to mitigate the lack of statistical parity in the output of the classifier is to relax the query upstream.

It is easy to construct additional examples that show how bias may be introduced during data cleaning, data integration, querying, and ranking — upstream from the final stage of data analysis. Therefore, it is meaningful to detect and mitigate these effects in the data lifecycle stages in which they occur. (See [Mitchell et al. 2018] for a discussion of the definitions of "bias", and of the corresponding assumptions made when defining fairness measures.)

Members of the data management community who are interested in this topic may consider a growing body of work on impossibility results, which show that different notions of fairness cannot be enforced simultaneously, and so require explicit trade-offs [Chouldechova 2017; Friedler et al. 2016; Kleinberg et al. 2017]. These are not negative results per se, nor are they surprising. Fairness is a subjective, context-dependent and highly politicized concept; a global consensus on what is fair is unlikely to emerge, in the context of algorithmic decision making or otherwise. Think, for example, of the decade-long debate about the interplay between "disparate treatment" and "disparate impact", for which recent examples include by Ricci v. De Stefano [3] and the ongoing lawsuit regarding the use of race in Harvard University admissions [4]. That being said, a productive way to move forward in the data science context is to develop methods that can be instrumented with different alternative fairness notions, and that can support principled and transparent trade-offs between these notions.

[2] https://www.fatconference.org/
[3] https://en.wikipedia.org/wiki/Ricci_v._DeStefano
[4] https://www.nytimes.com/2018/10/13/us/harvard-affirmative-action-asian-students.html

## 4 MOVING AND REMOVING PERSONAL DATA

### 4.1 The Right to Be Forgotten

The right to be forgotten is originally motivated by the desire of individuals to not be perpetually stigmatized by something they did in the past. Under pressure from despicable social phenomena such as revenge porn, it was turned recently into laws in 2006 in Argentina, and since then in the European Union, as part of the GDPR. In particular, Article 17 of the GDPR states that data subjects have the right to request erasure of their personal data, and that they can do so for a large number of reasons.

The passing of this law primarily resulted in a high number of requests to search engines to dereference web pages. This turned out to be controversial for a number of reasons, including also that the dereferencing by Google is very opaque, and that this company in effect acquired, against its own will, a questionable power to adjudicate. Furthermore, as is advocated by Wikimedia among others, the right to be forgotten sometimes conflicts with other rights such as the public's right to information.

In addition to search engines, the right to be forgotten affects companies that keep personal data. A prominent example is Facebook, where for many years it was impossible to delete data that pertains to a user's account. A user may close an account, then reopen it some time later and find all her data as it was originally. It is now possible to request the deletion of all data pertaining to an account from Facebook, however, the user has no proof that the deletion indeed occurred.

An important technical issue, of clear relevance to the data management community, is that of deletion of information in systems that are typically meant to accumulate data. This deletion must be both permanent and deep, in the sense that its effects must propagate through data dependencies. To start, it is difficult to guarantee that all copies of every piece of deleted data have actually been deleted. Further, when some data is deleted, the remaining database may become inconsistent, and may, for example, include dangling pointers. Additionally, production systems typically do not include a strong provenance mechanism, and so they have no means of tracking the use of an arbitrary data item (one to be deleted), and reasoning about the dependencies on that data item in derived data products.

Although much attention of the data management community has over the years been devoted to tracking and reasoning about provenance, primarily in relational contexts and in workflows (see [Herschel et al. 2017] for a recent survey), there is still important work to be done on making these methods both practically feasible, and sufficiently general to accommodate the current legal requirements. An important direction that is, to the best of our knowledge, still unexplored, concerns ascertaining the effects of a deletion on downstream processes that are not purely relational, but include other kinds of data analysis tasks, like data mining or predictive analytics.

Requests for deletion may also conflict with other laws such as requirements to keep certain transaction data for some period of time, or with requirements for fault tolerance and recoverability. Should the deleted pieces of data also be erased from caches and backups? Requesting this functionality gives immediate nightmares to systems engineers in charge of a production data management system, with millions of lines of code and terabytes of legacy data. The likely answer is: "this cannot be done; the only solution I see is redeveloping the system from scratch with right-to-be-forgotten-by-design." Understanding the impact of deletion requests on our ability to offer guarantees on system resilience and performance, and developing appropriate primitives and protocols for practical use, is another call to action for the data management community.

### 4.2 Interoperability and Portability

Article 20 of the GDPR, "Right to data portability", stipulates a data subject's right to receive her personal data from a vendor, and to transfer her data to another vendor. The main goals of this provision are both to keep the data subject informed about what data a vendor has about her, and to prevent vendor lock-in. This enables a user who is unhappy with a service to leave for a competing service that best serves her needs, without having to reconstruct her entire data history. This also allows a user to select applications of her choice and have them cooperate, to her best advantage, even if they come from different vendors.

In response to data portability regulation, and to users' concerns, Google, Twitter, Microsoft, and Facebook teamed up in the Data Transfer Project that aims to facilitate content transfer between applications. Of course, it is not an easy task for a company to provide a service that facilitates the departure of its customers. This is why, in spite of commendable behavior of companies that engage in the Data Transfer Project, it is the role of regulators to impose data portability and interoperability requirements.

Interoperability of database applications is an old topic. But one can imagine an unlimited number of possibilities, such as having a Whatsapp call talk to a Skype one. And it certainly acquires a different flavor when we consider interoperating applications with billions of users and millions of transactions per second.

For data portability, it should be noted that the devil is in the detail. The export format should be stable and structured to facilitate reuse. Also, which data can be exported is an issue. Obviously, it includes all data that the user volunteered to the service. But should it also include data the vendor gathered from the behavior of the user (e.g., the time the user is waking up in the morning)? Should it include data the service inferred (e.g., what is the home address of the user, her job address)?

Another issue with portability is the target system. A user may want to port her photos from Service A to Service B. The issue is then for Service B to be able to incorporate as much data as possible from Service A. Now, the user may want to integrate her photos in a personal information system [Abiteboul et al. 2015]. Such a system must be able to integrate information from a large panel of domains. This brings us to the fields of data integration [Lenzerini 2002] and knowledge representation.

## 5 NEUTRALITY

As already mentioned, Net Neutrality is now legally required in some countries. Yet, detecting Net Neutrality violations to enforce the law is not an easy task. Indeed, simply measuring the performance of Internet communications is not easy: measurement results may depend on the location of the source, of the target, of the context (other applications competing for the same bandwidth), and on other factors. Indeed, different measures provided for network traffic typically diverge. The evaluation of Net Neutrality relying on such hard-to-obtain measures is a challenging research topic [Molavi Kakhki et al. 2015], which is primarily of interest to the networks and Internet measurement communities, and less so to data management.

But beyond Net Neutrality, new forms of neutrality are emerging such as device neutrality (is my smart-phone blocking certain apps and favoring others?), and platform neutrality (is this particular web service providing neutral recommendation?). For instance, app stores like Google Play and the Apple App Store, tend to refuse to reference certain services, perhaps because they are competing with the company's own services. Research is needed to be able to verify these new facets of neutrality. In particular, it is not easy to check whether a recommendation engine like Google search or Booking is enforcing only transparent editorial policies, and whether, other than that, their results

are comprehensive, impartial and based solely on relevance. For example, it has been observed that search engines tend to favor some "friendly" services over competitors [5].

## 6 TAKE-AWAYS

In this article, we discussed several recent regulatory frameworks that aim to protect the rights of individuals, to ensure equitable treatment of services, and to bring transparency to data-driven algorithmic processes in industry and in government. Our goal was to bring these regulatory frameworks to the attention of the data management community, and to underscore the technical challenges they raise and which we, as a community, are well-equipped to address.

An important take-away of this article is that legal norms cannot be incorporated into data-driven systems as an afterthought. Rather, we must think in terms of *responsibility by design*, viewing it as a systems requirement.

We also stress that enacting algorithmic and data transparency, fairness, data protection, and neutrality will require a significant cultural shift. In making this shift, we must accept that the objectives of "efficiency", "accuracy" and "utility" cannot be the primary goal, but that they must be balanced with equitable treatment of members of historically disadvantaged groups, and with accountability and transparency to individuals affected by algorithmic decisions and to the general public.

In this article we focused on explicit regulation of industry stakeholders by government entities (in the case of the GDPR and the Net Neutrality laws), and on government oversight (in the case of the NYC ADS law). Another implicit regulatory mechanism can be achieved by empowering users and user associations, by providing them with data literacy education and with precise information on how different products and services work. Better educated users can choose better solutions, including more effective ways to protect their private data. Such users can also more easily understand explanations provided to them by an algorithmic system. User associations can help individuals make informed choices, and support them via class actions lawsuits in the case of disputes.

## ACKNOWLEDGMENTS

This work was supported in part by National Science Foundation (NSF) Grant No. 1741047, and by Agence Nationale de la Recherche (ANR) Grant Headwork.

## REFERENCES

Ziawasch Abedjan, Lukasz Golab, and Felix Naumann. 2017. Data Profiling: A Tutorial. In *Proceedings of the 2017 ACM International Conference on Management of Data, SIGMOD Conference 2017, Chicago, IL, USA, May 14-19, 2017*. 1747–1751. DOI: http://dx.doi.org/10.1145/3035918.3054772

Serge Abiteboul, Benjamin André, and Daniel Kaplan. 2015. Managing Your Digital Life. *Commun. ACM* 58, 5 (April 2015), 32–35. DOI: http://dx.doi.org/10.1145/2670528

Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2016. Machine Bias: Risk Assessments in Criminal Sentencing. *ProPublica* (May 23 2016). https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Alexandra Chouldechova. 2017. Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *CoRR* abs/1703.00056 (2017). http://arxiv.org/abs/1703.00056

Danielle K. Citron and Frank A. Pasquale. 2014. The Scored Society: Due Process for Automated Predictions. *Washington Law Review* 89 (2014). http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2376209

Anupam Datta, Shayak Sen, and Yair Zick. 2016. Algorithmic Transparency via Quantitative Input Influence: Theory and Experiments with Learning Systems. In *IEEE SP*. 598–617. DOI: http://dx.doi.org/10.1109/SP.2016.42

Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard S. Zemel. 2012. Fairness through awareness. In *Innovations in Theoretical Computer Science 2012, Cambridge, MA, USA, January 8-10, 2012*. 214–226. DOI: http://dx.doi.org/10.1145/2090236.2090255

[5] https://en.m.wikipedia.org/wiki/European_Union_vs._Google

Cynthia Dwork and Aaron Roth. 2014. The Algorithmic Foundations of Differential Privacy. *Foundations and Trends in Theoretical Computer Science* 9, 3-4 (2014), 211–407. DOI: http://dx.doi.org/10.1561/0400000042

Michael Feldman, Sorelle A. Friedler, John Moeller, Carlos Scheidegger, and Suresh Venkatasubramanian. 2015. Certifying and Removing Disparate Impact. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, NSW, Australia, August 10-13, 2015*. 259–268. DOI: http://dx.doi.org/10.1145/2783258.2783311

Sorelle A. Friedler, Carlos Scheidegger, and Suresh Venkatasubramanian. 2016. On the (im)possibility of fairness. *CoRR* abs/1609.07236 (2016). http://arxiv.org/abs/1609.07236

Government of India, Ministry of Communications. 2018. DoT Letter on Net Neutrality Regulatory Framework dated 31-07-2018. http://www.dot.gov.in/net-neutrality. (July 31 2018). [Online; accessed 29-November-2018].

Sara Hajian and Josep Domingo-Ferrer. 2013. A Methodology for Direct and Indirect Discrimination Prevention in Data Mining. *IEEE Trans. Knowl. Data Eng.* 25, 7 (2013), 1445–1459. DOI: http://dx.doi.org/10.1109/TKDE.2012.72

Melanie Herschel, Ralf Diestelkämper, and Houssem Ben Lahmar. 2017. A survey on provenance: What for? What form? What from? *VLDB J.* 26, 6 (2017), 881–906. DOI: http://dx.doi.org/10.1007/s00778-017-0486-1

H. V. Jagadish, Johannes Gehrke, Alexandros Labrinidis, Yannis Papakonstantinou, Jignesh M. Patel, Raghu Ramakrishnan, and Cyrus Shahabi. 2014. Big data and its technical challenges. *Commun. ACM* 57, 7 (2014), 86–94. DOI: http://dx.doi.org/10.1145/2611567

Faisal Kamiran, Indre Zliobaite, and Toon Calders. 2013. Quantifying explainable discrimination and removing illegal discrimination in automated decision making. *Knowl. Inf. Syst.* 35, 3 (2013), 613–644. DOI: http://dx.doi.org/10.1007/s10115-012-0584-8

Olga Kharif. September 2018. YouTube, Netflix Videos Found to Be Slowed by Wireless Carriers. *Bloomberg* (September 2018).

Keith Kirkpatrick. 2017. It's Not the Algorithm, It's the Data. *Commun. ACM* 60, 2 (Jan. 2017), 21–23. DOI: http://dx.doi.org/10.1145/3022181

Jon M. Kleinberg, Sendhil Mullainathan, and Manish Raghavan. 2017. Inherent Trade-Offs in the Fair Determination of Risk Scores. In *8th Innovations in Theoretical Computer Science Conference, ITCS 2017, January 9-11, 2017, Berkeley, CA, USA*. 43:1–43:23. DOI: http://dx.doi.org/10.4230/LIPIcs.ITCS.2017.43

Joshua A. Kroll, Joanna Huey, Solon Barocas, Edward W. Felten, Joel R. Reidenberg, David G. Robinson, and Harlan Yu. 2017. Accountable Algorithms. *University of Pennsylvania Law Review* 165 (2017). http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2765268

Maurizio Lenzerini. 2002. Data Integration: A Theoretical Perspective. In *Proceedings of the Twenty-first ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS '02)*. ACM, New York, NY, USA, 233–246. DOI: http://dx.doi.org/10.1145/543613.543644

Shira Mitchell, Eric Potash, and Solon Barocas. 2018. Prediction-Based Decisions and Fairness: A Catalogue of Choices, Assumptions, and Definitions. *CoRR* abs/1811.07867 (2018). https://arxiv.org/abs/1811.07867

Arash Molavi Kakhki, Abbas Razaghpanah, Anke Li, Hyungjoon Koo, Rajesh Golani, David Choffnes, Phillipa Gill, and Alan Mislove. 2015. Identifying Traffic Differentiation in Mobile Networks. In *Proceedings of the 2015 Internet Measurement Conference (IMC '15)*. ACM, New York, NY, USA, 239–251. DOI: http://dx.doi.org/10.1145/2815675.2815691

Haoyue Ping, Julia Stoyanovich, and Bill Howe. 2017. DataSynthesizer: Privacy-Preserving Synthetic Datasets. In *Proceedings of the 29th International Conference on Scientific and Statistical Database Management, Chicago, IL, USA, June 27-29, 2017*. 42:1–42:5. DOI: http://dx.doi.org/10.1145/3085504.3091117

Marco Túlio Ribeiro, Sameer Singh, and Carlos Guestrin. 2016. "Why Should I Trust You?": Explaining the Predictions of Any Classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016*. 1135–1144. DOI: http://dx.doi.org/10.1145/2939672.2939778

Andrea Romei and Salvatore Ruggieri. 2014. A multidisciplinary survey on discrimination analysis. *Knowledge Eng. Review* 29, 5 (2014), 582–638. DOI: http://dx.doi.org/10.1017/S0269888913000039

Julia Stoyanovich and Bill Howe. 2018. Follow the data! Algorithmic transparency starts with data transparency. https://ai.shorensteincenter.org/ideas/2018/11/26/follow-the-data-algorithmic-transparency-starts-with-data-transparency, *The Ethical Machine* (November 27 2018). [Online; accessed 19-March-2017].

Julia Stoyanovich, Bill Howe, Serge Abiteboul, Gerome Miklau, Arnaud Sahuguet, and Gerhard Weikum. 2017. Fides: Towards a Platform for Responsible Data Science. In *Proceedings of the 29th International Conference on Scientific and Statistical Database Management, Chicago, IL, USA, June 27-29, 2017*. 26:1–26:6. DOI: http://dx.doi.org/10.1145/3085504.3085530

The European Parliament AND Council. 2015. Regulation (EU) 2015/2120. https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:32015R2120. (2015). [Online; accessed on 28-September-2018].

The European Union. 2016. Regulation (EU) 2016/679: General Data Protection Regulation (GDPR). https://gdpr-info.eu/. (2016). [Online; accessed 28-September-2018].

The New York City Council. 2017. Int. No. 1696-A: A Local Law in relation to automated decision systems used by agencies. https://laws.council.nyc.gov/legislation/int-1696-2017/. (2017). [Online; accessed on 28-September-2018].

Tim Wu. 2003. Network Neutrality, Broadband Discrimination. *Journal of Telecommunications and High Technology Law* 2 (2003).

Ke Yang, Julia Stoyanovich, Abolfazl Asudeh, Bill Howe, H. V. Jagadish, and Gerome Miklau. 2018. A Nutritional Label for Rankings. In *Proceedings of the 2018 International Conference on Management of Data, SIGMOD Conference 2018, Houston, TX, USA, June 10-15, 2018*. 1773–1776. DOI: http://dx.doi.org/10.1145/3183713.3193568